

# An heuristic approach to obtain signal envelope with a simple software implementation

Cecilia Jarne<sup>a,b</sup>

<sup>a</sup>*Department of Science and Technology, University of Quilmes (UNQ) and CONICET*  
<sup>b</sup>*e-mail: cecilia.jarne@unq.edu.ar*

## Abstract

Signal amplitude envelope allows to obtain information of the signal features for different applications. It is widely used to pre-process sound and other signals of physiological origin in human or animal studies. In order to obtain signal envelope, a fast and simple algorithm is proposed based on peak detection. The procedure presented here is quite straightforward and can be used in different applications of time series analysis. It can be applied in signals with different origin and frequency content. This algorithm presented is implemented based on python libraries. An open source code is also provided. Aspects on the parameter selection are discussed to adapt the same method for different applications. Also traditional methods are revisited and compared with the one proposed here.

*Keywords:*

Algorithm, Signal analysis, Envelope, Rich spectral content, Python code, Open source

## 1. Introduction

In audio signals, envelope is usually defined together with temporal fine structure. The amplitude modulations of animal sounds, speech or musical instruments are important features that often can be used to understand the physics or biology behind different processes. Accurate estimation of the amplitude, or equivalently energy, or envelope of a time-domain signal (waveform) is not trivial. Ideally, the amplitude envelope should outline the waveform connecting the main peaks and avoiding over fitting as was pointed out in [1].

The definition of envelope in the literature is often ambiguous and lacks of an exact mathematical definition. It is commonly agreed that envelope

varies slowly and, in some empirical view, it should pass the prominent peaks of the data smoothly [2].

Amplitude envelope is a relevant acoustic feature. It allows to get a reliable estimation that follows closely sudden variations in amplitude and avoids ripples in more stable regions. Depending on the fundamental frequency of the signal, it is possible to have a near optimal order selection in some regions. Amplitude variations depending on the application allows to obtain information on the signal patterns or features. i. e. peak detection, maximum, signal sudden variations, etc.

In addition to sound signals, the signal envelope is useful in many application of signal analysis in different fields from biology or medicine, to geology or material sciences [3, 4, 5, 6, 7, 8, 9, 10]. A simple algorithm together with a code implementation could became a useful tool to be provide for experimental analysis across different disciplines.

Currently to obtain signal envelope there is a diversity of standard methods already implemented in different programming languages. Some of the most sophisticated has not been developed in open source code. On the other hand, most of the traditional methods do not allow to obtain a not attenuated envelope. Here a method to obtain the signal envelope is proposed and described and an open source implementation in python code is provided.

This work is organized as follows: Section 2 is used to provide a review of existing and most popular techniques. These techniques were also implemented with python libraries and compare with the algorithm proposed here. A comment on the more sophisticated algorithms is also presented, although no simple comparison with these techniques is possible due to the source the code is not open and available to perform it, the algorithms are not straightforward to implement and finally the analysis is beyond the scope of present work. The main purpose of this works is to propose a simple algorithm to get the amplitude signal envelope in any situation with a direct implementation. In Section 3.1 the algorithm proposed is described. The implementation is explained in Section 3.2 and some examples of application are provided in Section 3.3. This section includes the discussion of how to chose the parameters in order to adapt the method for different signals sources and some cues to lead to good practices in signal analysis are commented. The conclusion are presented in section 4. The code is available at the local University repository. In Section 4 the reference to download the code is provided.

## 2. Current envelope techniques

### 2.1. Classical techniques

The classical approach for the estimation of the amplitude envelope of a time-domain signal is the technique known as envelope follower [11, 12]. This method has been widely used and has several implementations across different programming languages and also analog circuits. It consists basically in rectifying the waveform and then low-pass filtering it. This can be implemented in the analog or digital domain. With programming tools, the algorithm consist of taking the absolute value of the digitalized signal, and then use a digital low-pass filter. Several programing languages has libraries with build-in filters such as MATLAB [13] or Python [14]. For the comparison that is presented in further in Section 3 we used python libraries.

Another interesting direct approach is to calculate the instantaneous root mean square (RMS) value of the waveform through a sliding window with finite support [15]. Some authors have proposed other techniques to obtain a more reliable estimation of the amplitude envelope of waveforms. An early attempt [12] consisted of a piece-wise linear approximation of the waveform. The amplitude envelope is created by finding and connecting the peaks of the waveform in a window that moves through the data. Again we used python libraries to generate a function for smooth the signal via RMS together with the sliding window and then compare it with the implementation of the method corresponding to the present work (Section 3).

Other classical technique consists of using the analytic signal from Hilbert Transform. Notably, if the Hilbert transform of  $S(t)$  is equal to its quadrature signal, then the estimates are equal to the actual information signals [11]. Synthetic signals (i.e. AM) can be constructed to have this property, but there is no reason to expect that animal sounds, acoustic musical instrument sounds or speech also present it. A more realistic condition is verified when we are dealing with narrow-band signals [16]. In sounds with rich spectral content it is not possible to use this approach because that condition is not fulfilled.

A straightforward approach is also an algorithm that directly use the prominent peaks and finally get the envelope by interpolation. A detail of this method can be found in [2]. It is an iterative process that needs to add noise to the original signal. In [2] not open source code is provide.

### 2.2. Regarding more sophisticated techniques

A sophisticated method developed with not attenuated envelope is called Cepstral Smoothing [1]. It uses the real cepstrum defined as the inverse

Fourier transform of the log magnitude spectrum. Regarding the log magnitude spectrum as a signal, it is possible to interpret each cepstral coefficient as a measure of the energy present in discrete frequency bands of that signal. Low-pass filtering the cepstrum would result in a smoother version of the log magnitude spectrum. If we only want to represent the spectral envelope, discarding information about the partials we should set the cutoff frequency below the period of the signal. This method is more sophisticated than the standard techniques and was discussed in detail in [1], but no software implementation has been provided.

Another method is called empirical mode decomposition (EMD) [17] allows to obtain the envelope as a result of an optimization process and sought as a minimum of a quadratic cost function. EMD is an adaptive data analysis method for analyzing nonlinear and non-stationary data, which was proposed by Huang et al. in 1998. A closed form solution of this optimization problem is obtained and it is shown that by choosing free parameters, where they can fine-tune the frequency resolution or the number of intrinsic mode functions (IMFs) as well as the shape of the envelopes, the method could be widely implemented.

Both methods showed to be interesting and accurate. Nevertheless as explained in Section 1 a disadvantage is that there are not published open source code implementations or libraries of these more sophisticated envelope techniques in any programming language of both works. This prevent us to perform the direct comparison of these algorithms with standard techniques or the one proposed in this paper without having to develop and debug additional code.

In following section the algorithm developed for present work is described.

### 3. Current method: an envelope based on a direct peak detection

#### 3.1. A three-step algorithm

To obtain an envelope output that will not be attenuated, a possible approach is to use a peak detection algorithm combined with a moving window. In the approach presented here, every value in the small window bunch is replaced by the maximum. Later it is possible to lowpass-filter the resulting signal afterwards to get rid of the remaining staircase. This process is summarized in the schematic presented in Figure 1. In more detail, the method implemented in present work consists of three simple steps. The steps are very basic and allows to obtain a smooth signal following the amplitude variations as precisely as required.

- **The first step** is to take absolute value of signal  $S(t)$ , meaning  $|S(t)|$  as it is shown in Figure 2 top panel.
- **In the second step** it is necessary to divide the  $|S(t)|$  in  $k$  in non overlapping bunches of  $N$  samples:

$$|S(t)| = S_1(t) + \dots + S_i(t) + \dots + S_k(t), \quad (1)$$

with each  $S_i(t) > 0$ . Then the maximum value is taken from the signal in each bunch:  $\text{Max}(|S_i(t)|) = M_i$ , corresponding to the  $j$  element value in each bunch (peak detection). At this point each value in the bunch is replaced by the maximum  $M_i$ . In this step signal ( $R(t)$ ) looks like Fig. 2 middle panel.

- **The third step** consist of a lowpass-filter applied to the resulted signal  $R(t)$  to get rid of the remaining staircase ripple to obtain  $\hat{S}(t)$  , as illustrated in the Figure 2 bottom panel.



Figure 1: Algorithm schema. Input signal is called  $S(t)$ . The envelope obtained is  $\hat{S}(t)$ . Each arrow represent one step of the algorithm.

The signal obtained with this technique has the same sample rate that the original signal. This could be very useful in order to compare the envelope with other simultaneous measurements related to the original signal.

By using this three-step method the envelope obtained follows the signal amplitude variations without attenuation, if the parameters are carefully selected as shown in Fig.2. In this example the filter cutoff frequency is chosen between  $100 - 150 \text{ Hz}$  to ensure that the signal envelope follows signal variations as rapid as  $10 \text{ mS}$  combined with the window bunch size of 35 samples. These two variables could be adjusted to follow variations as rapid as needed, depending on the required signal application.

A comparison between the classical implementations of envelope methods (the envelope follower and RMS approach) and the one proposed in this

work is presented in Figure 3. The code provided in section 4 allows to compared these methods for any audio signal of any sample rate and frequency content. In the example selected for the figure, the frequency cut for the envelope follower approach in this example is 150  $Hz$ . For the RMS mothod the slide window is 50 samples. For the three-step algorithm parameters are selected as in Figure 2.

### 3.2. Software implementation

The previous section was a general description of the method presented. Some particular considerations have to be taken into account for the code implementation and are discussed in this section. The libraries used, the function and parameters selected in the code are presented, as well as the arguments for the choices made.

The language used for the implementation was Python [14]. The main reason to use python is the expansive library of open source data analysis tools, also the web frameworks and the testing instruments that make one of the largest current programming community. Another reason is that Python is an accessible language for new programmers and the community provides many introductory resources.

Regarding the implementation parameters, the size of the bunch for the signal to be divided is an important one. It will mainly depend of the sampling rate  $f_S$  but also of the signal frequency content. For the examples presented here sample rate is 44.100  $KHz$  and frequency content is between 100  $Hz$  and 10000  $Hz$ . In this case a proper bunch size is between 20 – 200 samples. Since we have audio files available to analyze, the *Scipy Signal processing library* functions was used to get the audio sample rate and signal vector and write a code with these input variables.

With respect to the digital filter, it is necessary to choose a filter that does not produce phase shift differences (a zero-phase filtering). It means that does not shift the signal as it filters. Since the phase is zero at all frequencies, it is also linear-phase. Filtering backwards in time requires to predict the future signal value. This kind of filters can not be used for “online” real-life applications, only for off-line processing of recordings of signals or if we filter signal with some small-delay as it is acquired.

The *scipy signal library* [18] has two kinds of functions to chose for filter design. One is called *scipy.signal.lfilter* and the other *scipy.signal.filtfilt*. The first one (*lfilter*) is causal forward-in-time filtering only, similar to a real-life electronic filter. It cannot be zero-phase. It usually adds different

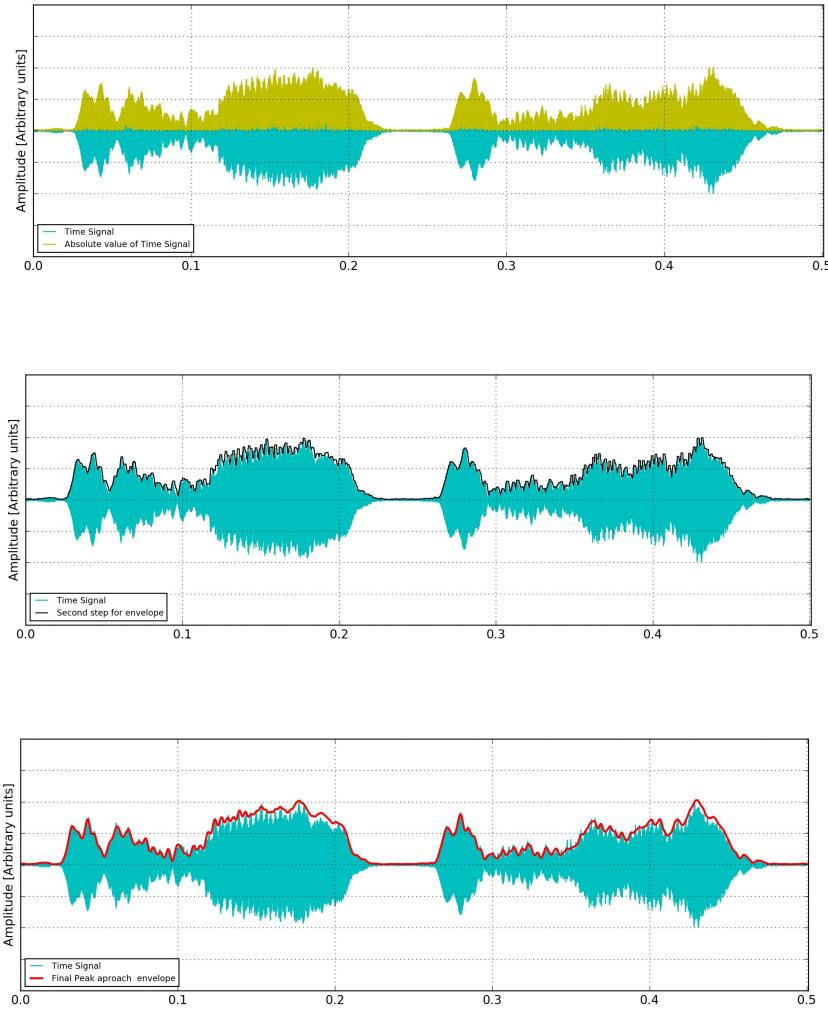


Figure 2: Zoom over a song canary segment. Top: Absolute value of time signal (First step). Middle: Peak detection algorithm (pre-envelope, second step). Bottom: envelope, the upper signal when a low pass filter is applied (third step). Filter cutoff frequency is 120 KHz and window bunch size of 35 samples.

amounts of delay at different frequencies. The appropriated method for the off-line application is to use *filtfilt* with no shift in the signal as it filters.

Once the zero-phase filter is selected, the following step consist of choosing the type of filter as well as the frequency behavior. In this case we used

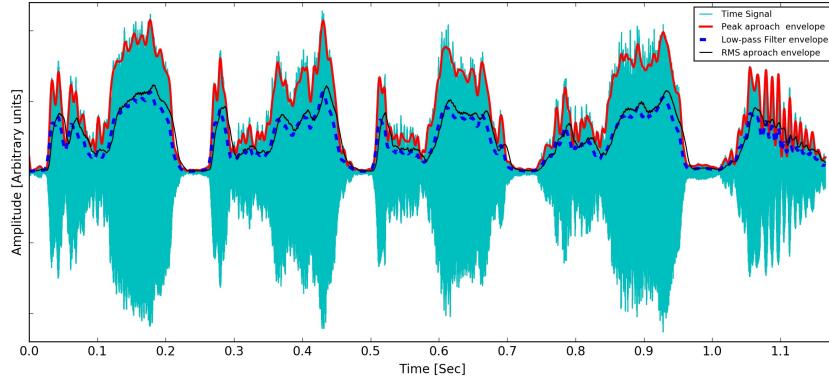


Figure 3: Comparison between the classical implementations of envelope methods and the one proposed in this work. Blue dashed line for low-pass filtered envelope, black thin line corresponds to RMS slide window and red line the proposed method.

a Butterworth low pass filter type `scipy.signal.butter` of fourth order. Frequency attenuation (behavior) of this filter is elsewhere [19]. The key point to obtain the envelope appropriate for the specific application is to choose the frequency cut of the low pass filter. Depending on this value you can obtain as many detail on the amplitude as it is required. Some examples for the same signal and different frequency cuts are presented in Section 3.3.

The method proposed here is not computational expensive, so it can be applied in different systems. For instance for 1.5 second of signal recording with a sample rate of 44.1 KHz, the computational time needed to run the code is < 0.5 Sec.

### 3.3. Example of the implementation over different sounds and parameter selection

A good method must be flexible enough to be applied to different kinds of signals. In this case the focus was on sound signal, but it is not exclusive. Four different sounds were selected from public data bases to test the algorithm. The envelope result from the method implemented is shown over different kinds of sound segments with or without rapid variations. Audio files was selected from public data sets: two sounds of very different animals presented in Figure 4 (canary song and a whale song), human speech (English language) and an instrument recording of piano song presented with corresponding spectrogram. Table 3.3 summarizes the parameter values for each case.

Time scales of signals are different as well as the frequency content and the parameter selection. A heuristic search for the values convenient in a particular application will depend on how many detail is required to follow the signal variation.

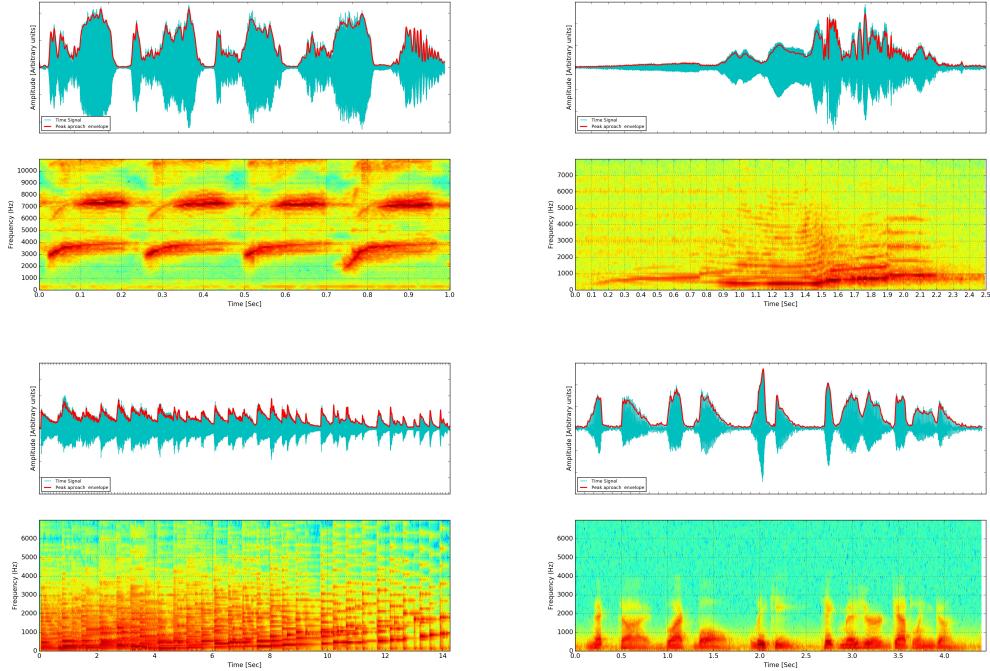


Figure 4: Up Left: Piano song segment and the corresponding sonogram. The frequency of the played notes increases over time. Up Right: Segment of spoken words in English. Down Left: A canary song signal with the estimated envelope and its sonogram. Down Right: A whale song with the corresponding envelope. Frequency content of both signals and also time scales are very different, although the method works fine in both cases.

Source	bunch size[samples]	Cut Frequency [Hz]
Canary sound [21]	35	300
Wheale sound [22]	50	300
Words spoken in english [23]	50	100
Piano stair scale notes [24]	200	100

Table 1: Examples for the values of the parameters selected for samples of different kind of sound origin. Sample rate is 44.1 *KHz* for all segments in the table.

#### 4. Conclusions

In this work a heuristic and simple technique is proposed for envelope estimation and a open source code is provided. This algorithm used is very simple to implement and it is not just another empirical method. Even when the procedure is straightforward and simple, its a simple and robust solution for different signals and required resolution on signal detail. An important advantage of this method is that it can be applied in signal with different spectral content. The time scaling parameter can be adjusted by the user in order to select the smoothness of the estimated envelope. Experimental applications presented shows that the algorithm can achieve good quality on the temporal envelope estimation of data. It was observed that the estimated envelope can track quite well the changing trend of input data in different time scales.

This algorithm can be implemented in off-line processing applications, e.g. animal sounds (particularly bird songs), musical instrument sound analysis, heart sound analysis, vibration signal analysis, and etc. In this paper, temporal audio signals were used as examples. However, there are no requirements on the specific domain of the data. It is appropriate for one dimensional data in any domain (e.g. time domain, frequency domain, space domain, etc.). Additionally a python implementation is presented to be used with any sound input.

A further key to apply the algorithm in almost real-time implementation, is to process signal time series in segments larger than 1 bunch size. It could be useful depending on the sample rate of the digitalized signal.

#### Source code

The software was developed on python. Code is available under password envelope at University local repository:

<http://ceciliajarne.web.unq.edu.ar/investigacion>

## Funding

This work was supported by the institutions: CONICET and UNQ.

## Acknowledgements

I want to thank to Gabriel Mindlin from *Laboratorio de Sistemas dinámicos* for advise and testing of the code. I want thank also the Stack Overflow Community for the general advice in the development of the code to be readable and easy to implement. A special mention for Cristina J. for her help in manuscript writing style. Finally I want to thank to Gabriel Lio for his personal help.

## References

- [1] M. Caetano, and X. Rodet. *Improved estimation of the amplitude envelope of time-domain signals using true envelope cepstral smoothing*. 2011 IEEE Int. Con. on Acoustics, Speech and Signal Processing (ICASSP)page 4244-4247.*DOI : 10.1109/ICASSP.2011.5947290*.
- [2] Q. Meng, M. Yuan, Z. Yang. *An empirical envelope estimation algorithm*. 2013 6th International Congress on Image and Signal Processing (CISP).*doi : 10.1109/CISP.2013.6745226*.
- [3] Michael J. Crosse, Edmund C. Lalor. *The cortical representation of the speech envelope is earlier for audiovisual speech than audio speech*. J Neurophysiol 111: 14001408, 2014; *doi : 10.1152/jn.00690.2013*.
- [4] S. Tilsen, A. Arvaniti. *Speech rhythm analysis with decomposition of the amplitude envelope: characterizing rhythmic patterns within and across languages*. The Journal of the Acoustical Society of America 134 (1), 628-639. *DOI : http://dx.doi.org/10.1121/1.4807565*.
- [5] M. Feldman. *Hilbert transform in vibration analysis*.Volume 25, Issue 3, April 2011, Pages 735802. *http://dx.doi.org/10.1016/j.ymssp.2010.07.018*.
- [6] A. Veltcheva1, P. Cavaco, C. Guedes Soares. *Comparison of methods for calculation of the wave envelope* Ocean Engineering Volume 30, Issue 7, May 2003, Pages 937948. *http://dx.doi.org/10.1016/S0029-8018(02)00069-0*.
- [7] J. Liu, W. Liu, H. Wang, T. Tao, J. Zhang. *A novel envelope extraction method for multichannel heart sounds signal detection*. International Proc. of Computer Science & Information Tech; 2012, Vol. 51, p630 *DOI : 10.7763/IPCSIT.2012.V51.106*.

- [8] Wu-bin Li, Chang-hou Lu, Jian-chuan Zhang. *A lower envelope Weber contrast detection algorithm for steelbar surface pit defects* Optics & Laser Technology 45 (2013) 654659. <http://dx.doi.org/10.1016/j.optlastec.2012.05.016>.
- [9] S. Choi, Z. Jiang. *Comparison of envelope extraction algorithms for cardiac sound signal segmentation*. Expert Systems with Applications 34 (2008) 10561069. doi : 10.1016/j.eswa.2006.12.015.
- [10] Wilsch A., Neuling T., Obleser J., Herrmann C.S. *Transcranial alternating current stimulation with speech envelopes modulates speech comprehension*. Neuroimage 2018. DOI : 10.1016/j.neuroimage.2018.01.038
- [11] A. Potamianos, P. Maragos. *A comparison of the energy operator and the Hilbert transform approach to signal and speech demodulation* Signal Processing 37 (1994) 95-120.
- [12] A. Schloss. *On the Automatic Transcription of Percussive Music - From Acoustic Signal to High-Level Analysis*, Stanford University, Ph.D. Thesis 1985.
- [13] MATLAB 6.1, The MathWorks Inc., Natick, MA, 2000
- [14] G. van Rossum. Python tutorial, Technical Report CS-R9526, Centrum voor Wiskunde en Informatica (CWI). <https://www.python.org/>. Amsterdam, May 1995.
- [15] J. Hajda. *A New Model for Segmenting the Envelope of Musical Signals: The relative Salience of Steady State versus Attack, Revisited*, J. AES, Nov. 1996.
- [16] W.M. Hartmann. *Signals, Sound, and Sensation*. Springer-Verlag-AIP Press, 1997.
- [17] Y. Washizawa, T. TanakaDanilo, P. M. Cichocki. *A Flexible Method for Envelope Estimation in Empirical Mode Decomposition*. Int. Con. on Knowledge-Based and Intelligent Information and Engineering Systems KES 2006. pp 1248-1255. DOI : 10.1007/11893011\_58.
- [18] Travis Oliphant, Pearu Peterson, Eric Jones 2001 *Scipy Library* <http://scipy.org/>
- [19] Butterworth digital and analog filter design library.  
[scipy-0.14.0/reference/generated/scipy.signal.butter.html](https://scipy-0.14.0/reference/generated/scipy.signal.butter.html)
- [20] Python Signal Processing Libraries.  
<https://docs.scipy.org/doc/scipy-0.14.0/reference/tutorial/signal.html>
- [21] Piano sound from: UCLA Phonetics Lab Language Archive  
<http://archive.phonetics.ucla.edu/Language/ENG/eng.html>

- [22] Canary sound from canary examples for training: <https://www.youtube.com/watch?v=qnZ-gj7Si2w&t=21s>
- [23] Whale sound from The Oceania Project Byron Bay, Australia. <https://soundcloud.com/iwhales>
- [24] Piano sound from Alexander Lerch <http://www.audiocontentanalysis.org/>